Title page

Establishing human leukemia xenograft mouse models by implanting human bone marrow-like scaffold-based niches

Short Title: Establishing a human leukemia xenograft mouse model

Antonella Antonelli[1], Willy A Noort[2], Jenny Jaques[1], Bauke de Boer[1], Regina de Jong-Korlaar[2], Annet Z Brouwers-Vos[1], Linda Lubbers-Aalders[2], Jeroen F van Velzen[3], Andries C Bloem[3], Huipin Yuan[4], Joost D de Bruijn[5], Gert J Ossenkoppele[2], Anton CM Martens[2], Edo Vellenga[1], Richard WJ Groen[2*], and Jan Jacob Schuringa[1*]

[1]Department of Experimental Hematology, Cancer Research Center Groningen (CRCG), University Medical Center Groningen, University of Groningen, The Netherlands. [2]Dept of Hematology, VU University Medical Center, Amsterdam, The Netherlands. [3]Dept of Immunology, University Medical Center Utrecht, Utrecht, The Netherlands. [4]Xpand Biotechnology BV, Bilthoven, The Netherlands. [5]Queen Mary University of London, School of Engineering and Materials Science (SEMS), Mile End Road, E1 4NS London, UK.

*** Corresponding authors**

Keywords: human leukemia xenograft mouse model; acute myeloid leukemia; leukemic stem cells; bone marrow microenvironment; AML

Full name, mailing address and email of corresponding authors: J.J. Schuringa, Department of Experimental Hematology, Cancer Research Center Groningen (CRCG), University Medical Center Groningen, University of Groningen, Hanzeplein 1, DA13, 9700RB, Groningen, The Netherlands. Phone: +31-503619391, fax: +31-503614862, email: j.j.schuringa@umcg.nl; and R.W.J. Groen, Department of Hematology, VU University Medical Center, De Boelelaan 1117, PK2 BR012, 1081HV, Amsterdam, The Netherlands. Phone: +31-204442345, fax: +31-204442601, email: r.groen@vumc.nl

Word count: 3993

Total number of figures: 6

Total number of Tables: 0

Key points

1. Humanized niche xenograft mouse models were generated that enable engraftment of leukemia patient cells covering all risk groups
2. Self-renewal is better maintained in the humanized niches as determined by serial transplantation and genome-wide transcriptome studies

Abstract

In order to begin to understand mechanisms that regulate self-renewal, differentiation and transformation of human hematopoietic stem cells, or to evaluate the efficacy of novel treatment modalities, stem cells need to be studied in their own species-specific microenvironment. By implanting ceramic scaffolds coated with human mesenchymal stromal cells (MSCs) into immune deficient mice we mimic the human bone marrow niche. Thus, we have established a human leukemia xenograft mouse model in which a large cohort of patient samples successfully engrafted covering all important genetic and risk subgroups. We find that by providing a humanized environment stem cell self-renewal properties are better maintained as determined by serial transplantation assays and genome-wide transcriptome studies, and less clonal drift is observed as determined by exome sequencing. The human leukemia xenograft mouse models that we have established here will serve as an excellent resource for future studies aimed at exploring novel therapeutic approaches.

Introduction

In order to study the molecular mechanisms involved in human leukemias and to improve treatment options, establishing *in vivo* xenograft models that faithfully recapitulate the disease is essential. Currently, NOD-SCID IL2Rγ$^{-/-}$ (NSG) xenograft models are considered to be the gold standard for evaluating engraftment of human haematological malignancies. However, this model has serious drawbacks since 30-40% of primary AML samples fail to engraft even in this most immune-compromised NSG model.[1-3] In addition, engraftment of normal human CD34^{+} cells in NSG mice is lymphoid biased and *in vivo* myeloid transformation has been much more difficult to achieve in NOD-SCID based xenograft models, including the NSG mice.[4-8] Furthermore, using retro/lentiviral systems to model human leukemias we frequently observe that while *in vitro* both myeloid and lymphoid transformation can be achieved upon expression of oncogenes such as MLL-AF9 or BCR-ABL/BMI1, *in vivo* transformation in mice is strongly biased towards ALL.[5,6] This is in contrast to patients expressing the BCR-ABL p210 oncoprotein, which can give rise to both myeloid and lymphoid leukemias, and the same is true for MLL-AF9 in paediatric leukemia patients. Together, these observations suggest that a specific human microenvironment might be necessary for engraftment and maintenance of self-renewal of myeloid leukemic stem cells.

Since certain myeloid growth factors are often species-specific, the murine bone marrow niche is most likely not sufficient to provide the appropriate environment for human LSCs. NOD/SCID-3/GM/SF mice engineered to produce human IL3, GM-CSF and Steel Factor/SCF[9-11] and NOD-SCID IL2Rγ$^{-/-}$ (NSG) mice expressing human IL3, GM-CSF and SCF (NS-SGM3)[3] have been generated, and enhanced engraftability of primary human AML samples was observed in these mouse strains. Also, the expression of these three human growth factors was sufficient to allow AML development upon transplantation of CB CD34^{+} cells expressing MLL-AF9.[3] Various other

mouse strains have also been developed[11-17] including MISTRG mice expressing M-CSF, IL3, GM-CSF, TPO and SIRPα[12] and are therefore useful tools, although other human niche-specific factors might still be missing.

To develop *in vivo* mouse leukemia xenograft models that would more faithfully recapitulate human leukemias we made use of a humanized niche xenograft model in which ceramic scaffolds coated with human mesenchymal stromal cells (MSCs) used to mimic the bone marrow microenvironment are implanted in immunodeficient mice.[18,19] Next we studied the engraftment of a large cohort of primary AML patient samples covering all important genetic and risk subgroups. Our data indicate that the presence of a humanized microenvironment favors engraftment of AML patient samples, including favorable risk AML samples that are notoriously difficult to engraft in NSG strains and that self-renewal is better maintained in the human niche.

Methods

Establishment of the humanized scaffold niche xenograft model and transplantations

The ectopic bone model was established as described previously.[18,19] Briefly, four to six hybrid scaffolds consisting of three 2–3mm biphasic calcium phosphate (BCP) particles loaded with human MSCs were implanted subcutaneously into female NOD.Cγ-Prkdcscid Il2rγtm1Wjl/SzJ (NSG) or RAG2$^{-/-}$γc$^{-/-}$ mice. Eight weeks after scaffold implantation, different cell doses ranging from 5x10^4 to 4x10^6 were directly injected into the scaffolds as indicated during primary as well as secondary transplantations.

Patient samples

Peripheral blood and bone marrow from untreated patients diagnosed with AML/ BC-CML/ B-ALL at the UMCG or VUmc were studied after informed consent and protocol approval by the Medical Ethical Committee of the UMCG or VUmc, in accordance with the Declaration of Helsinki. After ficoll separation of mononuclear cells, cells were cryopreserved until further use. $CD34^+$ cells were enriched or $CD3^+$ cells were depleted using a magnetically activated cell-sorting CD34 progenitor kit or automatically by using auto Macs (Miltenyi Biotech) as described previously.[6,20]

Additional Methods are described in the Supplemental Materials

Results

Establishing xenograft human leukemia models using three-dimensional scaffolds coated with human MSCs.

A cohort of 39 patients was selected to establish humanized niche mouse xenograft leukemia models (Supplemental Table 1; 17/39 (44%) adverse risk group, 8/39 (21%) intermediate risk and 14/39 (36%) favorable risk). Cells were injected into a humanized model that is based on subcutaneous implantation of human BM-like scaffolds (huBM-sc) in order to provide a human microenvironment, as described previously.[18,19] Briefly, in this model ceramic scaffolds coated with human mesenchymal stromal cells (MSCs) were implanted subcutaneous in immunodeficient mice where they developed into structures mimicking a human bone marrow microenvironment including bone formation, and ossicles were vascularized by murine blood vessels (Figure 1B, Supplemental Figure 1, Supplemental Figure 2A). Six to eight weeks later, $CD34^+$ or $CD3^+$-

depleted AML mononuclear cells (MNCs) were then injected in these so-formed "ossicles". As comparator, 6 of the patient samples were also injected intravenously (i.v., referred to as muBM-iv) in mice without human scaffolds to evaluate engraftment in a murine microenvironment (Figure 1A, Supplemental Table 1). As second comparator, cells were injected either directly in the scaffold or i.v. in mice that did carry humanized scaffolds in 5 selected cases (referred to as IV huBM-sc, Figure 1A, Figure 1E). Experiments were carried out independently in two institutes using 2 different mouse strains, NSG (at the UMCG, NL) and $Rag2^{-/-}\gamma c^{-/-}$ mice (at the VU, NL) (Supplemental Table 1) showing robustness of the model. Leukemic cell engraftment was monitored over time with PB sampling and/or by monitoring tumor volume in the scaffolds (Figure 1A, Supplemental Table 2). Hematoxylin and Eosin (HE) staining of non-injected scaffold sections 6 weeks after implantation confirmed the presence of extramedullary bone in the scaffolds. Different magnifications reveal the presence of stromal cells and bone material deposited in the cavities of the ceramic particle, and blood vessels of mouse origin were also observed (Figure 1B, Supplemental Figure 1A, Supplemental Figure 2A).

For huBM-sc mice, chimerism levels in the PB were usually low (Supplemental Table 2), so the volume of the biggest tumor growing on one of the scaffolds at approximately 1.5-2 cm^3 was defined as endpoint of the experiment. MuBM-iv mice were terminated when we observed $huCD45^+$ levels in the PB around 30-60% of total living cells together with signs of illness. We obtained engraftment and outgrowth of leukemic cells in 3 out of 7 cases with the muBM-iv model (43%) and in 29 out of 39 cases (74%) in the huBM-sc model. Lack of engraftment was not due to an absence of bone formation in the scaffolds (Supplemental Figure 1B). Notably, favorable risk leukemias only engrafted in the huBM-sc model but no engraftment was observed in any of the muBM-iv mice without humanized scaffolds, and the same was true for patient sample #2 that also

did not engraft in the muBM-iv model (Figure 1C, Supplemental Table 1). All 7 favorable risk AMLs with an inv16, which are notoriously difficult to engraft in NSG mice, efficiently engrafted in the huBM-sc mice (Supplemental Table 1). AML #1 engrafted in both the huBM-sc as well the muBM-iv models with similar kinetics of leukemia development, while AMLs #3 and #4 did engraft in the muBM-iv model but with slower kinetics as compared to the huBM-sc model (Supplemental Table 2).

All intermediate risk samples engrafted in both models (Figure 1C, Figure 1D). Interestingly, in huBM-sc mice leukemic cells engrafted not only in the humanized scaffold niches but also seeded into murine hematopoietic organs, with engraftment levels in murine compartments increasing concomitantly with prognosis of disease (e.g. the percentage of human leukocytes in muBM ranged from 4 to 20% in the favorable risk mice while chimerism levels in murine BM, spleen and liver were significantly higher in mice injected with adverse risk samples reaching up to an average of 55% in one case of adverse risk leukemia) (Figure 1D). Five AML samples were i.v. injected in mice carrying humanized scaffolds (Figure 1A, route 3). All samples efficiently seeded into the scaffolds while only in 2/5 cases engraftment was also observed in the human bone marrow (Figure 1E).

The onset of tumor initiation and consequently the time of sacrifice (14-38 weeks) were variable from patient to patient, with tumor growth rates that appeared to correlate with risk group. In the majority of animals, tumors be palpable 100 days after injection with the exception of 2 cases of adverse risk AML (#4 and #2) where the onset was earlier than 100 days (Figure 1F, Supplemental Table 2). Furthermore, using a luciferase gene-marked AML sample, we observed that growth was initiated 50 days after engraftment (Supplemental Figure 2C). In all cases of engrafted intermediate

and adverse risk leukemias, the growth rate (read out by the slope of the curve) was faster than in favorable risk AMLs (Figure 1F, Supplemental Table 2).

Leukemic cells engrafted in huBM-sc mouse models recapitulate the original phenotype of the patient.

FACS analysis was performed to compare the immunophenotype of the leukemic cells of the original patient samples with tumor cells isolated from the human scaffold niche as well as from the murine compartments. In some cases the original patient immunophenotype was well conserved *in vivo* regardless of whether the cells expanded in a humanized or murine microenvironment, while in other cases the original patient immunophenotype was clearly better preserved in the humanized scaffolds. An example of the latter is AML #2. This sample contained an inv16 and a t(9;22) BCR-ABL translocation with no mutations in NPM or FLT3. At presentation, this patient displayed a relatively large immature population of blast cells (34% $CD34^+$, 48% $CD117^+$ $CD15^-$) and all MNCs were positive for CD33 and CD13, weakly positive for CD14, CD11b, CD11c and CD15, and negative for CD19 (Figure 2A). $CD34^+$ blast cells were isolated from this patient and injected in the scaffolds of NSG mice and within 14 weeks tumors developed on the scaffolds and the animals were terminated (Figure 2B, Supplemental Table 2). Scaffold tumors were greenish in color (indicating the presence of myeloid myeloperoxidase expressing cells) and were well vascularized with murine blood vessels. Leukemic cells could also be retrieved from murine BM, spleen, liver and PB (Supplemental Table 2). Cells retrieved from the humanized scaffolds were $huCD45^+$ with percentages of $CD34^+$ and $CD117^+$ comparable to the percentages at diagnosis, whereas in cells retrieved from BM the percentage of $CD34^+$ and in particular $CD117^+$ cells was partially reduced while CD15 and to some extent also CD11b and

CD14 percentages were increased (Figure 2C, Figure 2D). Histological analysis using HE staining confirmed that scaffold tumors contained bone and huCD45^{neg} stromal cells surrounded by huCD45 and huCD33-positive AML blasts (Figure 2E). MGG staining of cytospins confirmed that scaffold tumors mainly contained immature blast-like cells with big nuclei and BM samples revealed a more abundant presence of smaller cells with a more differentiated morphology (Figure 2F).

In the case of AML #1, the humanized niche did not clearly confer an advantage to preserve a more immature phenotype compared to the murine stromal environment (Supplemental Figure 3). This sample contained a FLT3-ITD, was NPMwt and contained several chromosomal abnormalities including a t(3;5) NPM-MLF1 translocation, with a high percentage of CD34^{+}/CD117^{+} blasts (Supplemental Figure 3A). Within 100-150 days tumors were visible and again infiltration of human CD45^{+} cells was observed in mouse organs as well (Supplemental Figure 3B-C, Supplemental Table 2). However, the immunophenotype of the original patient sample #1 was now conserved in both the humanized niche as well as in mouse niches (Supplemental Figure 3C-F).

Leukemic cells isolated from humanized scaffolds have superior 2nd transplantation capacity compared to leukemic cells isolated from the murine bone marrow niche.

Secondary transplantation assays were performed for three cases using various doses ranging from 4.5x10^{4}-3x10^{6} cells harvested from scaffolds or bone marrow that were injected into individual scaffolds of secondary mice (Figure 3A). Secondary engraftment transplantation into 2nd scaffolds could readily be established for huBM-sc-harvested cells in all investigated cases (Figure 3A-G and Supplemental Figure 2D). For AML #2 secondary transplantation could readily be established for huBM-sc-harvested cells with 3x10^{6}, 1x10^{6} and even 4x10^{5} cells, but no 2nd engraftment was

observed for muBM-iv-harvested cells using up to $4x10^5$ cells (Figure 3B-C, Supplemental Table 2). The same was true for AML #3 whereby 2nd engraftment was observed with $1.5x10^6$ and $1x10^5$ injected huBM-sc-harvested cells, but not with up to $4x10^5$ muBM-iv-harvested cells (Figure 3D-E, Supplemental Table 2). For AML #1, 2nd engraftment could be established with as little as $4.5x10^4$ huBM-sc-harvested cells, which was not seen with muBM-iv-harvested cells when injected at the same dose, but injection of $1.8x10^5$ muBM-iv-harvested cells did allow secondary transplantation, albeit with slower kinetics as compared to huBM-sc-harvested cells at the same dose (Figure 3F-G, Supplemental Table 2). These data appeared to be in line with flow cytometry analyses on stem/progenitor compartments suggesting that the immature $CD34^+/CD38^-$ compartments were better preserved in the human niches (Supplemental Figure 4A). Together, these data again pinpoint to the notion that stemness was better preserved in the humanized niche for most leukemias, whereby AML #1 seemed to be the least dependent on a human microenvironment within the tested cohort, in line with our observation that this patient sample was also able to engraft in the muBM-iv model, while e.g. AML #2 appeared to be much more dependent on cues from a human microenvironment. We also analyzed lymphoid engraftment, but did not find any $CD3^+/CD4^+$ or $CD3^+/CD8^+$ cells (Supplemental Figure 4B).

Next, we compared engraftment of a favorable risk, intermediate risk and adverse risk sample in mice carrying scaffolds coated with either human MSCs or murine MSCs and performed primary and secondary engraftment experiments (Figure 4A). The adverse risk sample engrafted in both conditions, regardless of whether scaffolds were coated with human or murine MSCs (Figure 4B). The intermediate risk sample showed the highest chimerism levels in scaffolds coated with human MSCs compared to murine MSCs, and secondary engraftment was only observed with cells harvested from human scaffolds (Figure 4B). Finally, the favorable risk sample efficiently

engrafted in the human scaffolds, and only at very late timepoints in the murine scaffolds (day 322) (Figure 4B). Efficient secondary transplantation was observed in all three cases in mice carrying scaffolds coated with human MSCs. No significant differences were seen in secondary transplantation of adverse risk sample #27 when scaffolds coated with human versus murine MSCs were compared, but significantly delayed secondary engraftment was observed with intermediate risk sample #18 in the scaffolds coated with murine MSCs. Failure of engraftment in the murine scaffolds was not due to insufficient formation of an extramedullary murine niche development (Figure 4D).

Evaluation of clonal heterogeneity and clonal drift within humanized and murine niches.

We selected three cases (#1, #2 and #3) in order to evaluate clonal heterogeneity and clonal drift. Variant allelic frequencies (VAF, exome seq) of recurrent mutations were determined in patient samples at presentation, and in tumor material harvested from the humanized scaffold niche, from the murine BM niche, and from *in vitro* expanded AML on MS5 bone marrow stromal cells supplemented with human cytokines. For #2 the $CD34^+$ fraction of the patient, two huBM-sc samples, one muBM sample and *in vitro* expanded sample could be analyzed. As shown in Figure 5A, a number of mutations were detected with approximately 50% VAF, including XBP1, APC, RPTOR, FLNC and MLL2 mutations, which were detected at the start, in the huBM-sc samples as well as in the in vitro expanded sample. In all these cases, the inv16 and t(9;22) translocation were detected as well (data not shown). In contrast, these mutations were not detected in the muBM expanded sample, which had lost the inv16, t(9;22) and mutations indicated above, but instead displayed IDH1 mutations with an VAF of 50% which was only detected at low percentage in the diagnostics sample, and various other mutations that had remained below the detection limit in the

diagnostic sample (Figure 5A), clearly indicating a change in clonal composition in the absence of a humanized microenviroment. For AML #1 this clonal drift was not seen, although a small clone or clones appeared with mutations in e.g. BCOR, KDM6A, PTPN11 and APC in the muBM that was not detected at diagnosis or in the huBM-sc samples, while at day 43 in the *in vitro* expanded cells again a different clone(s) appeared with mutations in HDAC7, IKZF1, MLL and XBP1 (Figure 5B). In sample #3 the main clone(s) with approximately 50% VAF in NRAS, MSH6, DNMT3A was present at diagnosis and in the huBM-sc as well as the muBM samples. The diagnosis sample also contained mutations in MUTYH and PLCG2 with a VAF of approximately 50% that was also seen in the huBM-sc sample but only at a very low VAF in the muBM sample suggesting that this was an independent clones that did not efficiently grow out in the murine niche. Finally, new small clone(s) appeared as well with low VAF in both the huBM-sc and muBM samples that were not detected in the diagnostic sample (Figure 5C). Together, these data clearly indicate that clonal heterogeneity and clonal drift are important parameters to evaluate in *in vitro* and *in vivo* models, and that a humanized niche might be better suited to maintain clonal heterogeneity observed at diagnosis.

Human MSCs are better in supporting long-term expansion of human AML compared to murine bone marrow stroma.

In vitro experiments were performed using either human MSCs, murine MS5 bone marrow stromal cells or liquid culture without stroma using #2 and #1 as representative examples. In all cases, $CD34^+$-sorted AML cells were grown with or without a cocktail of human cytokines IL3, G-CSF and TPO as we previously described.[21-23] No long-term cultures could be established without stroma, although some expansion was observed for #2 in liquid-culture conditions with cytokines,

but within 20 days cells differentiated and stopped expanding (Supplemental Figure 5A). Human MSCs clearly were superior in supporting long-term expansion of both AML samples even in the absence of additional cytokines (Supplemental Figure 5A). The addition of cytokines did further enhance expansion in both murine MS5 as well as in human MSC cocultures, but the immature $CD117^{+}$ phenotype was clearly better preserved in the absence of cytokines, in particular for AML #2 and to a lesser degree for AML #1 (Supplemental Figure 5B-C).

Stemness transcriptome signatures are better preserved in a humanized niche.

Transcriptome studies were performed on diagnostic samples as well as on huBM-sc and muBM derived samples from leukemic mice for 6 patient samples (Supplemental Table 3). Unsupervised hierarchical cluster analysis revealed that each patient displayed a unique transcriptome signature regardless of where cells were retrieved from (huBM-sc, muBM or diagnostic sample) (Figure 6A). Supervised clustering within individual patient samples revealed that diagnostic samples clustered most closely together with the huBM-sc samples for AML cases #2, #3, #6 and #4 (Figure 6B). This was not the case for AML #1 for which we had indeed already observed that this sample was less dependent on a human environment for its self-renewal and stemness properties, and also for the B-ALL sample #5 (Figure 6B). Gene Ontology (GO) analyses revealed that specific gene sets were differently regulated in the human versus murine niche, including processes such as transcription regulation, immune response, inflammatory response, mitochondria, and regulation of apoptosis (Figure 6C). In line with our secondary transplantation studies, GSEA analyses revealed significant enrichment for Leukemic Stem Cell signatures for #2 and #3 when grown in the humanized scaffold compared to the murine BM, while this was not observed for sample #1 (Figure 6D). Instead, a significant enrichment for leukemic-GMP signatures was observed (Figure

6D). Examples of gene expression of individual genes are shown in Figure 6E, indicating that expression of stem cell-related genes such as BMI1, HOXA5, HOXA9, CD34, CDKN1C and GATA2 was better preserved in the human niche compared to the murine niche for AML #2, while this was much less explicit for AML #1.

Discussion

Hematopoietic stem cells, and the same holds true for leukemic stem cells, do not simply live alone. They are surrounded by a variety of other cell types that together constitute the stem cell niche in the bone marrow.[24-28] In order to begin to understand mechanisms that regulate self-renewal, differentiation and transformation of human hematopoietic stem cells, both intrinsic mechanisms as well as extrinsic mechanisms involving cues from the environment need to be taken into account.[29,30] Even though the currently available immune deficient NSG xenograft mouse models are considered the gold standard for evaluating engraftment of human haematological malignancies, these models have serious drawbacks since only a limited percentage of primary AML patient samples can engraft and these models are strongly lymphoid biased.[1-8] Clearly, a human bone marrow microenvironment that would provide a suitable home for normal and leukemic stem cells is lacking in these animals. To capture and maintain stem cell self-renewal programs and clonal heterogeneity as is observed in patients, a human niche appears to be essential. By making use of scaffolds coated with human MSCs to create a humanized environment in mice we have established human xenograft mouse leukemia models that can serve as a resource for leukemic stem cell studies and the evaluation of novel therapeutic approaches. We find that by providing a humanized environment: i) 29 out of 39 (74%) leukemia cases could efficiently engraft in the huBM-sc model covering all important genetic subtypes and risk groups, whereby also all

favorable risk inv(16) patient samples could engraft, which are notoriously difficult to engraft in NSG mice; ii) stem cell self-renewal is better preserved in essentially all investigated cases in the huBM-sc as compared to the mouse bone marrow niche as determined by serial transplantation assays; iii) clonal heterogeneity as observed in the original patient is much better preserved in the huBM-sc compared to the mouse BM, at least in one case; and iv) stem cell self-renewal signatures are better preserved in the huBM-sc in the majority of cases. Even when cells were injected i.v. in mice carrying human scaffolds, better engraftment was observed in all 5 studied cases, whereby in 3/5 cases no or hardly any engraftment was observed at all in the mouse bone marrow compartment. When scaffolds were coated with mouse MSCs instead of human MSCs, a favorable risk sample failed to engraft, and an intermediate risk sample engrafted less efficiently in primary mice, and not at all in secondary mice. During the course of our studies, Reinisch et al also described that ossicles generated by injecting human MSCs with matrigel in xenograft mice could generate a humanized niche in which leukamia patient samples could engraft more efficiently, including favorable risk PML-RARa samples that are notoriously difficult to engraft in normal NSG mice.[31] These data are nicely in line with ours providing independent confirmation of the robustness of these humanized niche xenograft mouse models, but we also provide evidence that stem cell self-renewal is better preserved in a humanized niche as determined by transcriptome studies and serial transplantations.

Obviously, heterogeneity exists between leukemia patients, and we find that some patient samples are more dependent on the presence of a human niche than others. As examples, we extensively studied 2 cases: AML #2 which strongly depended on a humanized environment, and AML #1 for which this was much less the case. Consistently, AML #2 could not engraft when i.v.-injected in mice without human scaffolds, the original patient phenotype defined by flow and transcriptome

signatures defined by genome-wide transcriptome studies was better maintained in the human scaffolds, self-renewal properties were better preserved in the human scaffolds, and finally the dominant BCR-ABL/inv(16) clone that was present in the original patient sample also grew out in humanized scaffolds while in the mouse bone marrow this clone was absent and instead an minor clone carrying e.g. an IDH1 mutation expanded. Similar observations were published by others[32], showing that upon transplantation of AML samples in NSG or even NSG-SGM3 mice clonal drift can occur, stressing the need for careful genomic analysis of engrafted clones. At the other end of the spectrum, AML #1 could also engraft when injected i.v. into mice without human scaffolds, the original patient phenotype defined by flow and transcriptome signatures was more comparable between human and murine niches, and secondary engraftment could be achieved with muBM-harvested cells as well, albeit with lower frequencies as compared to huBM scaffold-retrieved cells. Overall, our data indicate that the humanized scaffold models will serve as an excellent resource for future studies aimed at exploring novel therapeutic approaches. Experiments were carried out independently in 2 institutes using 2 different mouse strains, NSG and RAG2$^{-/-}\gamma c^{-/-}$ mice, showing the robustness of the model. Clinical trials often fail when evaluating drugs that appeared successful in *in vitro* and *in vivo* studies[33,34], and in part this might be related to the models used, whereby a humanized environment was not included. Finally, the current model allows for an efficient evaluation of drug efficacy as well, as already documented for multiple myeloma [18] and MLL-AF9 models[35], indicating that the human leukemia xenograft mouse models that we have established here will serve as an excellent resource for future studies aimed at exploring novel therapeutic approaches.

Acknowledgements

This work was supported by grants from the European Research Council (ERC-2011-StG 281474-huLSCtargeting) to JJS. RG was supported by a grant from the Dutch Cancer Foundation (VU2011-5127). Dr. G.J. Schuurhuis is acknowledged for providing clinical data of a subset of AML specimen. The authors would like to thank dr. G. Vanasse (Novartis, USA) for help with exome sequencing.

Author Contributions

Conceptual design of experiments JJS and RWJG; mouse studies at the UMCG (Groningen, NL) were performed by AA, JJ and JJS; mouse studies at the VU (Amsterdam, NL) were performed by WAN, RdJK, LLA, ACMM, and RWJG; JFvV and ACB analyzed FACS-data; transcriptome studies were performed and analyzed by AZBV and JJS; scaffold material was provided by HY and JDB; AA, BdB, WAN, GJO, EV, ACMM, RWJG and JJS analyzed data; AA, RWJG and JJS wrote the manuscript.

Disclosure of Conflicts of Interest: The authors declare no conflicts of interest.

References

Reference List

1. Vargaftig J, Taussig DC, Griessinger E, et al. Frequency of leukemic initiating cells does not depend on the xenotransplantation model used. *Leukemia*. 2011;26(4):858-860.

2. Sanchez PV, Perry RL, Sarry JE, et al. A robust xenotransplantation model for acute myeloid leukemia. *Leukemia*. 2009;23(11):2109-2117.

3. Wunderlich M, Chou FS, Link KA, et al. AML xenograft efficiency is significantly improved in NOD/SCID-IL2RG mice constitutively expressing human SCF, GM-CSF and IL-3. *Leukemia*. 2010;24(10):1785-1788.

4. Barabe F, Kennedy JA, Hope KJ, Dick JE. Modeling the initiation and progression of human acute leukemia in mice. *Science*. 2007;316(5824):600-604.

5. Horton SJ, Jaques J, Woolthuis C, et al. MLL-AF9-mediated immortalization of human hematopoietic cells along different lineages changes during ontogeny. *Leukemia*. 2012;51116-1126.

6. Rizo A, Horton SJ, Olthof S, et al. BMI1 collaborates with BCR-ABL in leukemic transformation of human CD34+ cells. *Blood*. 2010;116(22):4621-4630.

7. Wei J, Wunderlich M, Fox C, et al. Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. *Cancer Cell*. 2008;13(6):483-495.

8. Townsend EC, Murakami MA, Christodoulou A, et al. The Public Repository of Xenografts Enables Discovery and Randomized Phase II-like Trials in Mice. *Cancer Cell*. 2016;29(4):574-586.

9. Nicolini FE, Cashman JD, Hogge DE, Humphries RK, Eaves CJ. NOD/SCID mice engineered to express human IL-3, GM-CSF and Steel factor constitutively mobilize engrafted human progenitors and compromise human stem cell regeneration. *Leukemia*. 2004;18(2):341-347.

10. Feuring-Buske M, Gerhard B, Cashman J, Humphries RK, Eaves CJ, Hogge DE. Improved engraftment of human acute myeloid leukemia progenitor cells in beta 2-microglobulin-deficient NOD/SCID mice and in NOD/SCID mice transgenic for human growth factors. *Leukemia*. 2003;17(4):760-763.

11. Goyama S, Wunderlich M, Mulloy JC. Xenograft models for normal and malignant stem cells. *Blood*. 2015.

12. Rongvaux A, Willinger T, Martinek J, et al. Development and function of human innate immune cells in a humanized mouse model. *Nat.Biotechnol.* 2014;32(4):364-372.

13. Rongvaux A, Willinger T, Takizawa H, et al. Human thrombopoietin knockin mice efficiently support human hematopoiesis in vivo. *Proc.Natl.Acad.Sci.U.S.A.* 2011;108(6):2378-2383.

14. Strowig T, Rongvaux A, Rathinam C, et al. Transgenic expression of human signal regulatory protein alpha in Rag2-/-gamma(c)-/- mice improves engraftment of human hematopoietic cells in humanized mice. *Proc.Natl.Acad.Sci.U.S.A.* 2011;108(32):13218-13223.

15. Willinger T, Rongvaux A, Takizawa H, et al. Human IL-3/GM-CSF knock-in mice support human alveolar macrophage development and human immune responses in the lung. *Proc.Natl.Acad.Sci.U.S.A.* 2011;108(6):2390-2395.

16. Willinger T, Rongvaux A, Strowig T, Manz MG, Flavell RA. Improving human hemato-lymphoid-system mice by cytokine knock-in gene replacement. *Trends Immunol.* 2011;32(7):321-327.

17. Theocharides AP, Rongvaux A, Fritsch K, Flavell RA, Manz MG. Humanized hemato-lymphoid system mice. *Haematologica*. 2016;101(1):5-19.

18. Groen RW, Noort WA, Raymakers RA, et al. Reconstructing the human hematopoietic niche in immunodeficient mice: opportunities for studying primary multiple myeloma. *Blood*. 2012;120(3):e9-e16.

19. Gutierrez A, Pan L, Groen RW, et al. Phenothiazines induce PP2A-mediated apoptosis in T cell acute lymphoblastic leukemia. *J.Clin.Invest*. 2014;124(2):644-655.

20. Schuringa JJ, Chung KY, Morrone G, Moore MA. Constitutive activation of STAT5A promotes human hematopoietic stem cell self-renewal and erythroid differentiation. *J.Exp.Med.* 2004;200(5):623-635.

21. Schuringa JJ and Schepers H. Ex vivo assays to study self-renewal and long-term expansion of genetically modified primary human acute myeloid leukemia stem cells. *Methods Mol.Biol.* 2009;538287-300.

22. Sontakke P, Carretta M, Capala M, Schepers H, Schuringa JJ. Ex vivo assays to study self-renewal, long-term expansion, and leukemic transformation of genetically modified human hematopoietic and patient-derived leukemic stem cells. *Methods Mol.Biol.* 2014;1185195-210.

23. van Gosliga D, Schepers H, Rizo A, van der Kolk D, Vellenga E, Schuringa JJ. Establishing long-term cultures with self-renewing acute myeloid leukemia stem/progenitor cells. *Exp.Hematol.* 2007;35(10):1538-1549.

24. Calvi LM and Link DC. The hematopoietic stem cell niche in homeostasis and disease. *Blood*. 2015;126(22):2443-2451.

25. Calvi LM, Adams GB, Weibrecht KW, et al. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature*. 2003;425(6960):841-846.

26. Zhang J, Niu C, Ye L, et al. Identification of the haematopoietic stem cell niche and control of the niche size. *Nature*. 2003;425(6960):836-841.

27. Mendez-Ferrer S, Michurina TV, Ferraro F, et al. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature*. 2010;466(7308):829-834.

28. Schepers K, Campbell TB, Passegue E. Normal and leukemic stem cell niches: insights and therapeutic opportunities. *Cell Stem Cell*. 2015;16(3):254-267.

29. Enver T, Pera M, Peterson C, Andrews PW. Stem cell states, fates, and the rules of attraction. *Cell Stem Cell*. 2009;4(5):387-397.

30. Eaves CJ. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood*. 2015;125(17):2605-2613.

31. Reinisch A, Thomas D, Corces MR, et al. A humanized bone marrow ossicle xenotransplantation model enables improved engraftment of healthy and leukemic human hematopoietic cells. *Nat.Med.* 2016.

32. Klco JM, Spencer DH, Miller CA, et al. Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell*. 2014;25(3):379-392.

33. Alizadeh AA, Aranda V, Bardelli A, et al. Toward understanding and exploiting tumor heterogeneity. *Nat.Med.* 2015;21(8):846-853.

34. Gould SE, Junttila MR, de Sauvage FJ. Translational value of mouse models in oncology drug development. *Nat.Med.* 2015;21(5):431-439.

35. Sontakke P, Carretta M, Jaques J, et al. Modeling BCR-ABL and MLL-AF9 leukemia in a human bone marrow-like scaffold based xenograft model. *Leukemia*. 2016;in press.

Figures Legends

Figure 1. Overview of the mouse xenograft leukemia models. A, Schematic representation of the generation of muBM-iv and huBM-sc models. **B,** Representative HE stain of scaffold sections 6 weeks after implantation; b, bone; s, scaffold; v, blood vessels. **C,** Success rate of leukemia development in muBM-iv and huBM-sc models, expressed as percentage of AML samples from

different risk groups, which engrafted or not in the two xenograft models. **D**, Scatter plots showing engraftment of donor human $CD45^+$ cells in different compartments of the mouse at sacrifice (huBM-sc, humanized scaffolds; muBM, murine bone marrow; muSp, murine spleen; muLv, murine liver; muPB, murine peripheral blood). Engraftment values from several mice for each AML sample are expressed as mean±SEM. Route 1: cells were directly injected into the humanized scaffolds, route 2: cells were injected i.v. in mice without humanized scaffolds. **E**, Experiment as in (d), but now cells were injected i.v. in mice carrying humanized scaffolds. **F**, *In vivo* tumor growth rates of the biggest tumor scaffold of each mice. Each colored line represents one mouse from a certain AML sample. Tumor size (cm^3) was used as endpoint of the experimental period.

Figure 2. Mouse xenograft AML model using sample AML #2. A, FACS immunophenotype of AML #2 patient-derived MNCs cells. **B,** Representative photograph depicting the "leukemic masses" formed around the extramedullary bones of Sc1 and Sc3 and two scaffolds with no tumor growth (Sc4, not injected control scaffold), at mouse sacrifice. **C**, Representative FACS phenotype from a primary huBM-sc mouse (m23, see Table S1) transplanted with AML #2 cells. Expression of a panel of hematopoietic markers in $huCD45^+$ gated cells isolated from different scaffolds or from the murine BM. **D**, Bar graph summarizing the FACS phenotype of human $CD45^+$ cells from different compartments of all mice injected with #2 AML cells, compared with original patient phenotype. Values are expressed as mean±SEM. **E**, Histologic sections of scaffold tumor from a representative #2 huBM-sc mouse (m24, Sc1) stained with HE, anti-huCD45 or anti-huCD33 antibodies. **F**, Representative images (m23) of May Grünwald Giemsa (MGG)-stained cytospins of cells retrieved from Sc1, Sc3 or BM.

Figure 3. Assessment of the self-renewal potential of engrafting AML cells by serial transplantation of three AML samples. A, Setup of secondary transplantation experiments. HuCD45$^+$ cells harvested from from a scaffold or from the BM of primary huBM-sc mice were injected in the scaffolds of secondary recipient NSG mice implanted with the humanized niches 6-8 weeks before the transplantation of leukemic cells. **B,D,F**, Representative FACS analysis of secondary recipient mice of AML #2, AML #3 and AML #1, respectively. **C,E,G**, *In vivo* tumor growth rates of the biggest tumor scaffold in secondary transplanted mice from AML #2, AML #3 and AML #1, respectively . Each colored line represents one mouse and the cell dosage and source (SC=human scaffold or BM=mouse bone marrow) are depicted in the corresponding legend. Tumor size (cm^3) was used as endpoint of the experimental period.

Figure 4. Assessment of the self-renewal potential of engrafting AML cells by serial transplantation of three AML samples. A, Schematic representation of the generation of mouse models with scaffolds coated with human MSCs (huBM-sc) versus murine MSCs (muBM-sc). **B**, Plots showing % engraftment of donor human CD45$^+$ cells in the scaffolds in huBM-sc versus muBM-sc models, using patient samples with adverse, intermediate and favorable risk. Engraftment values from 8 scaffolds in 2 mice (mouse 1 indicated with red symbols and mouse 2 indicated with black symbols) for each AML sample are expressed as mean±SEM. **C**, For secondary transplantations, leukemic cells harvested from the scaffold of a primary huBM-sc or muBM-sc mouse were injected in the scaffolds of secondary recipient huBM-sc or muBM-sc mice. Kaplan Meijer plots are shown. Significant better secondary engraftment for the intermediate risk sample #18 was shown in the huBM-sc model, while significant better engraftment for the favorable risk sample (#33) was already observed in the primary transplantations. **D**, Ceramic

scaffold implants coated with murine MSCs generate an ectopic hematopoietic niche. HE stains of different murine implants, before inocculation with AML. As shown the scaffolds (s) are covered by murine bone (b) with vascularization (v), resulting in the support of the murine hematopoiesis in Rag2-/-γc-/- mice.

Figure 5. *In vivo* and *in vitro* maintenance of the genetic heterogeneity of the original AML patient samples. Exome sequencing to determine variant allele frequency in the leukemic populations for 3 representative AML samples: **A**, #2, **B**, #1, **C**, #3. Vertical axis shows the percentage of variant reads per allele, corresponding to different somatic mutations indicated in the legend. On the horizontal axis, several samples from same AML are depicted. All samples were sorted for huCD45 expression before sequencing. VAF around 50% indicated clonal mutations with heterozygous configuration. VAF around 100% would indicate clonal mutations with homozigous configuration, whilst frequencies lower than 50% show the presence of subclonal mutations.

Figure 6. Gene expression profiling of patient derived-AML cells and engrafted in the humanized or murine hematopoietic niche of huBM-sc mice. A, Unsupervised hierarchical cluster analysis (Pearson uncentered absolute distance, average linkage) on gene expression profiles of muBM- or huSC-engrafted huCD45^{+} cells or patient- derived (MNCs or CD34^{+}) leukemic cells from six different AMLs. **B**, Supervised cluster with Euclidean distance and ward linkage analyses per AML using genes that were differentially expressed between huBM-sc and muBM-retrieved cells (fold change (FC) ≥ 3). **C**, GO analyses on genes that were differentially expressed between huBM-sc and muBM-retrieved cells (fold change (FC) ≥ 3). **D**, GSEA analyses

on genes that were differentially expressed between huBM-sc and muBM-retrieved cells (fold change (FC) ≥ 3). **E**, Expression of individual stem cell-related genes in the diagnostic patient sample and huBM-sc and muBM-retrieved cells.

Figures

Fig. 1

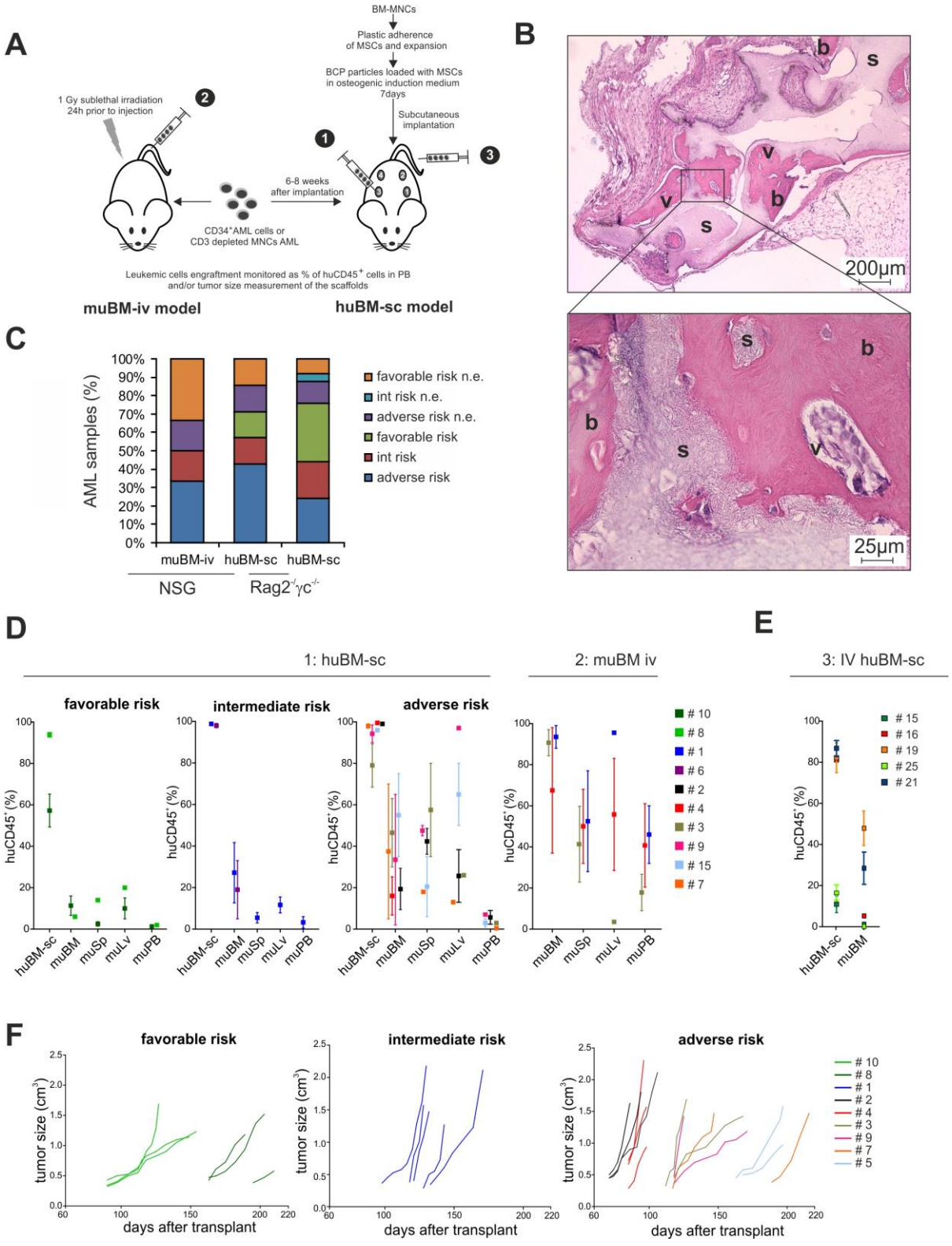


Fig. 2

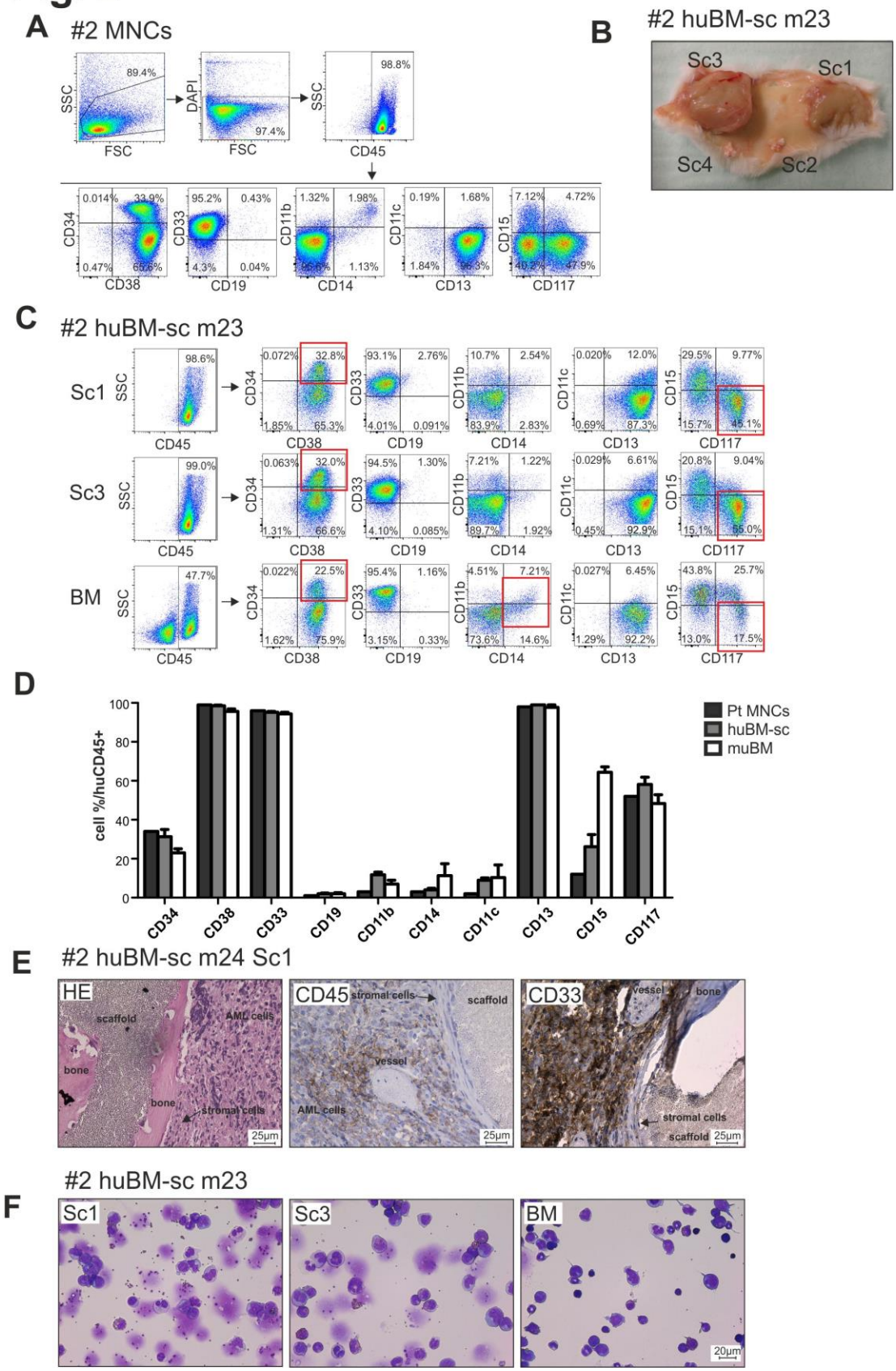
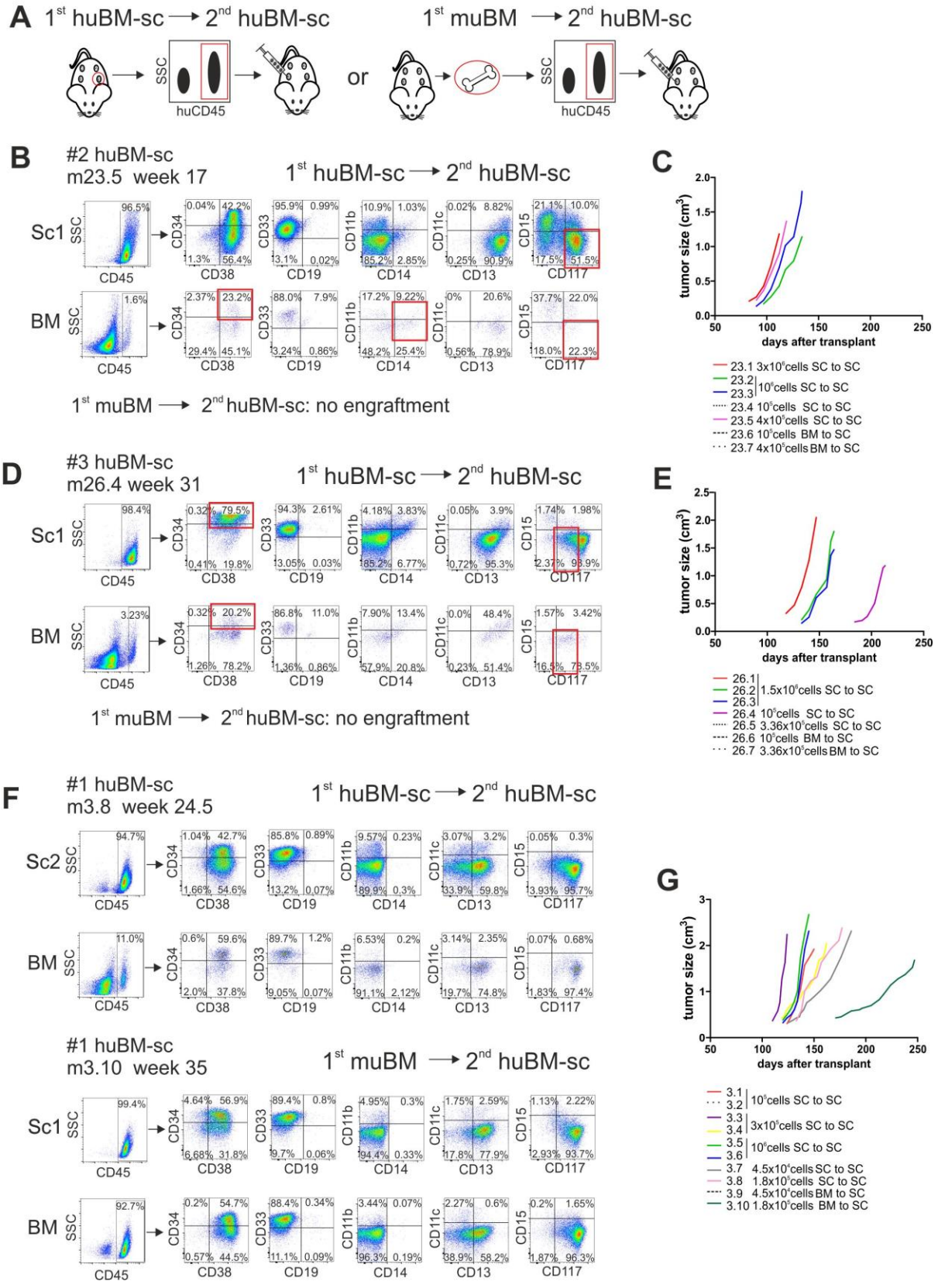
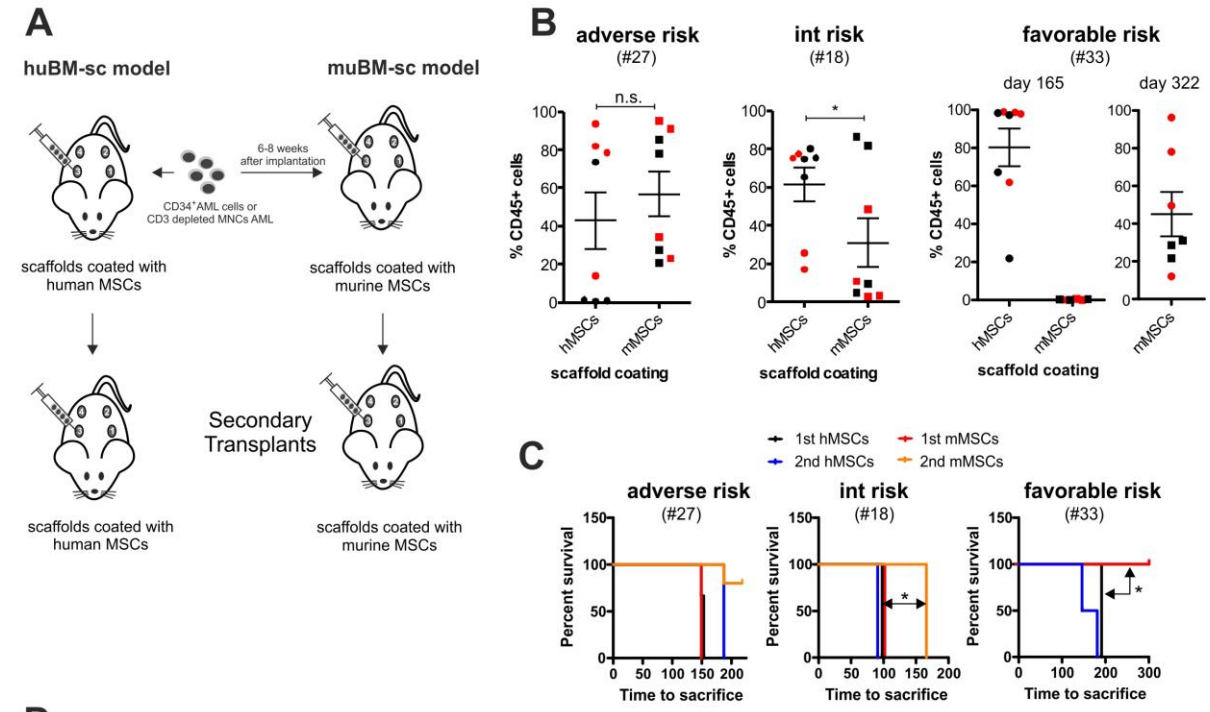

A
#2 MNCs
SSC
FSC
89.4%
DAPI
97.4%
98.8%
CD45
CD34
CD38
CD33
CD19
CD11b
CD14
CD11c
CD13
CD15
CD117
B
#2 huBM-sc m23
Sc3
Sc1
Sc4
Sc2
C
#2 huBM-sc m23
Sc1
Sc3
BM
D
cell %/huCD45+
Pt MNCs
huBM-sc
muBM
E
#2 huBM-sc m24 Sc1
HE
scaffold
AML cells
bone
stromal cells
CD45
vessel
CD33
25μm
F
#2 huBM-sc m23
Sc1
Sc3
BM
20μm

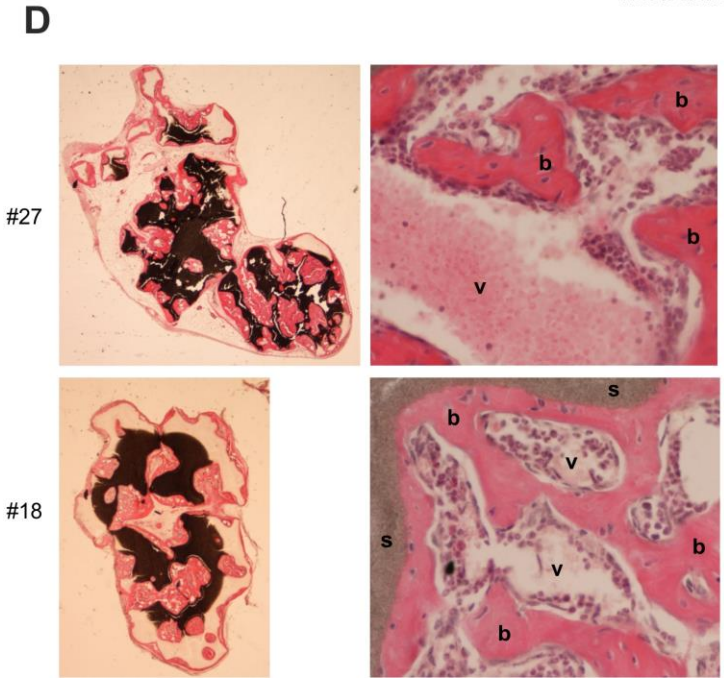
Fig. 3


A
1st huBM-sc → 2nd huBM-sc
1st muBM → 2nd huBM-sc
or
SSC
huCD45
B
#2 huBM-sc
m23.5 week 17
1st huBM-sc → 2nd huBM-sc
Sc1
BM
SSC
CD45
CD34
CD38
CD33
CD19
CD11b
CD14
CD11c
CD13
CD15
CD117
1st muBM → 2nd huBM-sc: no engraftment
C
tumor size (cm3)
days after transplant
23.1 3x10^6cells SC to SC
23.2 10^6cells SC to SC
23.3 10^6cells SC to SC
23.4 10^5cells SC to SC
23.5 4x10^5cells SC to SC
23.6 10^5cells BM to SC
23.7 4x10^5cells BM to SC
D
#3 huBM-sc
m26.4 week 31
1st huBM-sc → 2nd huBM-sc
1st muBM → 2nd huBM-sc: no engraftment
E
26.1 26.2 26.3 1.5x10^6cells SC to SC
26.4 10^5cells SC to SC
26.5 3.36x10^5cells SC to SC
26.6 10^5cells BM to SC
26.7 3.36x10^5cells BM to SC
F
#1 huBM-sc
m3.8 week 24.5
1st huBM-sc → 2nd huBM-sc
Sc2
#1 huBM-sc
m3.10 week 35
1st muBM → 2nd huBM-sc
G
3.1 3.2 10^6cells SC to SC
3.3 3.4 3x10^6cells SC to SC
3.5 3.6 10^6cells SC to SC
3.7 4.5x10^4cells SC to SC
3.8 1.8x10^5cells SC to SC
3.9 4.5x10^5cells BM to SC
3.10 1.8x10^5cells BM to SC

1

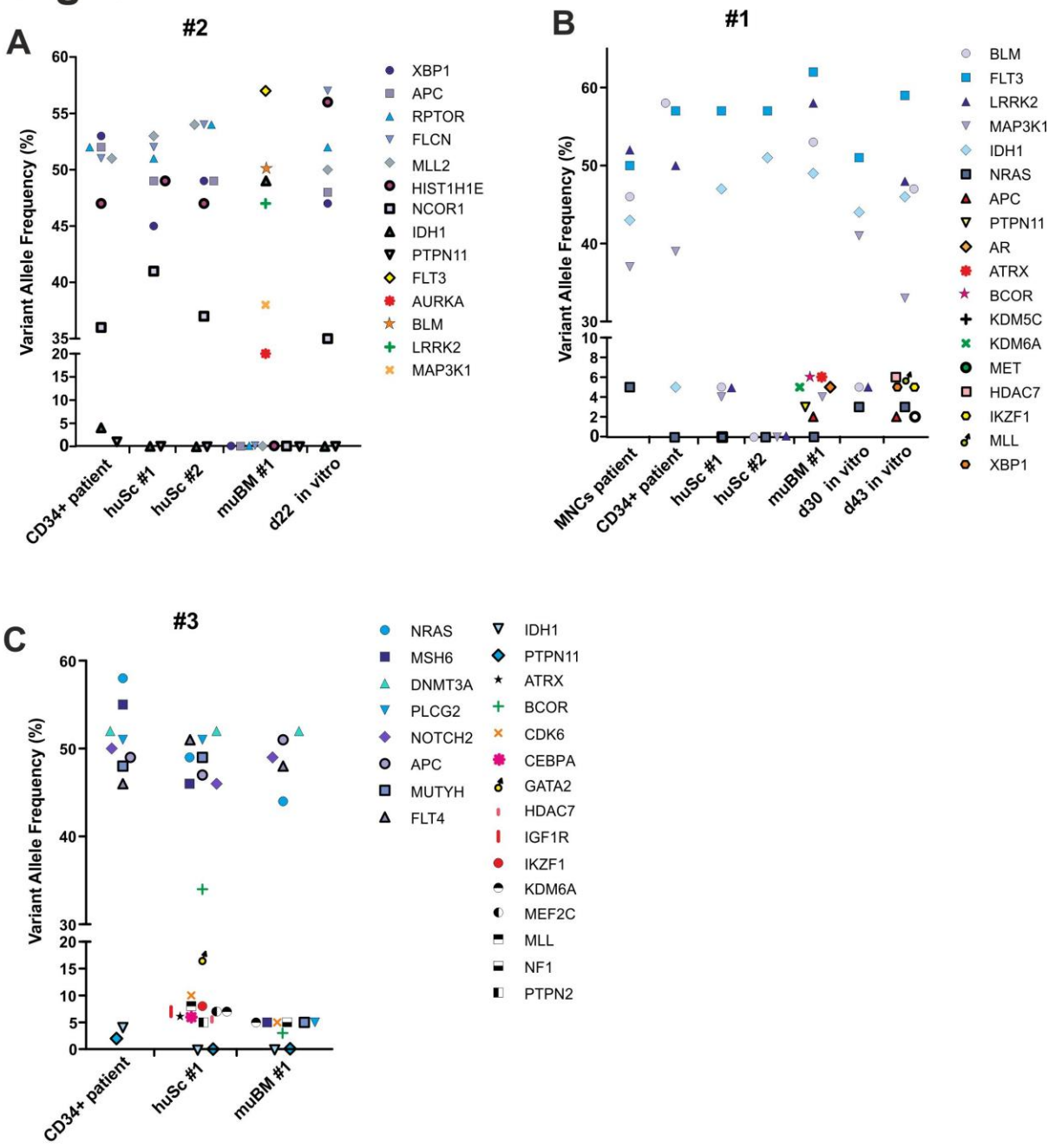
Fig. 4


A
huBM-sc model
muBM-sc model
6-8 weeks after implantation
CD34+AML cells or CD3 depleted MNCs AML
scaffolds coated with human MSCs
scaffolds coated with murine MSCs
Secondary Transplants
scaffolds coated with human MSCs
scaffolds coated with murine MSCs
B
adverse risk (#27)
int risk (#18)
favorable risk (#33)
day 165
day 322
n.s.
*
% CD45+ cells
hMSCs
mMSCs
scaffold coating
C
1st hMSCs
1st mMSCs
2nd hMSCs
2nd mMSCs
adverse risk (#27)
int risk (#18)
favorable risk (#33)
Percent survival
Time to sacrifice


D
#27
#18
b
v
s

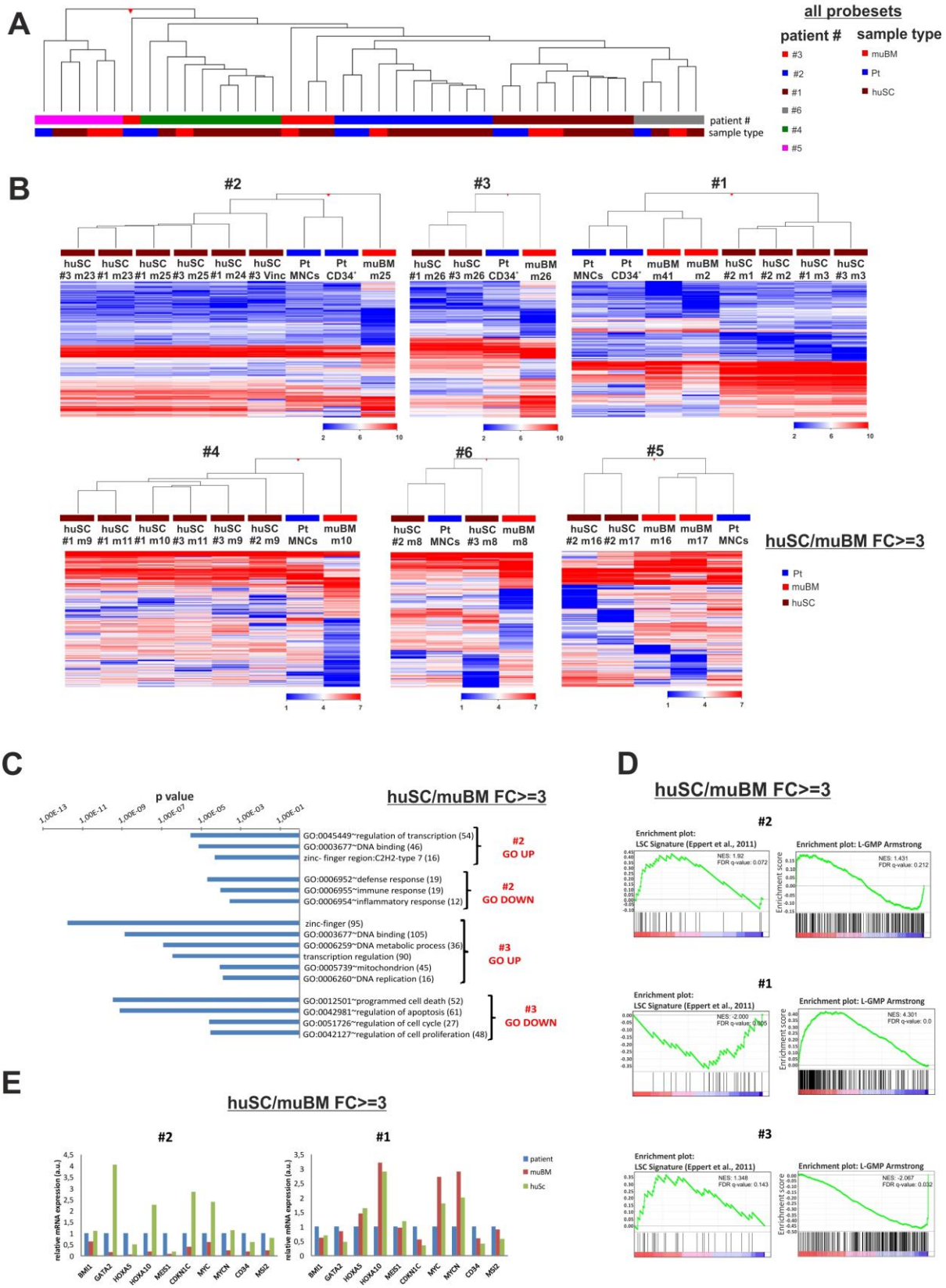
1

Fig. 5


A
#2
Variant Allele Frequency (%)
CD34+ patient
huSc #1
huSc #2
muBM #1
d22 in vitro
XBP1
APC
RPTOR
FLCN
MLL2
HIST1H1E
NCOR1
IDH1
PTPN11
FLT3
AURKA
BLM
LRRK2
MAP3K1
B
#1
Variant Allele Frequency (%)
MNCs patient
CD34+ patient
huSc #1
huSc #2
muBM #1
d30 in vitro
d43 in vitro
BLM
FLT3
LRRK2
MAP3K1
IDH1
NRAS
APC
PTPN11
AR
ATRX
BCOR
KDM5C
KDM6A
MET
HDAC7
IKZF1
MLL
XBP1
C
#3
Variant Allele Frequency (%)
CD34+ patient
huSc #1
muBM #1
NRAS
MSH6
DNMT3A
PLCG2
NOTCH2
APC
MUTYH
FLT4
IDH1
PTPN11
ATRX
BCOR
CDK6
CEBPA
GATA2
HDAC7
IGF1R
IKZF1
KDM6A
MEF2C
MLL
NF1
PTPN2

1

Fig.6


A
all probesets
patient #
sample type
#3
#2
#1
#6
#4
#5
muBM
Pt
huSC
patient #
sample type
B
#2
#3
#1
#4
#6
#5
huSC/muBM FC>=3
Pt
muBM
huSC
C
p value
huSC/muBM FC>=3
GO:0045449~regulation of transcription (54)
GO:0003677~DNA binding (46)
zinc- finger region:C2H2-type 7 (16)
#2 GO UP
GO:0006952~defense response (19)
GO:0006955~immune response (19)
GO:0006954~inflammatory response (12)
#2 GO DOWN
zinc-finger (95)
GO:0003677~DNA binding (105)
GO:0006259~DNA metabolic process (36)
transcription regulation (90)
GO:0005739~mitochondrion (45)
GO:0006260~DNA replication (16)
#3 GO UP
GO:0012501~programmed cell death (52)
GO:0042981~regulation of apoptosis (61)
GO:0051726~regulation of cell cycle (27)
GO:0042127~regulation of cell proliferation (48)
#3 GO DOWN
D
huSC/muBM FC>=3
#2
#1
#3
Enrichment plot:
LSC Signature (Eppert et al., 2011)
Enrichment plot: L-GMP Armstrong
E
huSC/muBM FC>=3
#2
#1
relative mRNA expression (a.u.)
patient
muBM
huSc
BMI1
GATA2
HOXA5
HOXA10
MEIS1
CDKN1C
MYC
MYCN
CD34
MSI2